Communicative and social consequences of interactions with voice assistants

Aeschlimann, Sara ; Bleiker, Marco ; Wechner, Michael ; Gampe, Anja

Abstract: The growing prevalence of artificial intelligence and digital media in children's lives provides them with the opportunity to interact with novel non-human agents such as robots and voice assistants. Previous studies show that children eagerly adopt and interact with these technologies, but we have only limited evidence of children's distinction between artificial intelligence and humans. In this study, the communication patterns and prosocial outcomes of interactions with voice assistants were investigated. Children between 5 and 6 years (N = 72) of age solved a treasure hunt in either a human or voice assistant condition. During the treasure hunt, the interaction partner supplied information either about their knowledge of or experience with the objects. Afterwards, children were administered a sharing task and a helping task. Results revealed that children provided voice assistants with less information than humans and that only the type of information given by a human interaction partner was related to children's information selection. Sharing was influenced by an interaction between type of information and interaction partner, showing that the type of information shared influenced children's decisions more when interacting with a human, but less when interacting with a voice assistant. Children in all conditions enjoyed the treasure hunt with the interaction partner. Overall, these results suggest that children do not impose the same expectations on voice assistants as they do on humans. Consequently, cooperation between humans and cooperation between humans and computers differ.

DOI: https://doi.org/10.1016/j.chb.2020.106466

Computers in Human Behavior 112 (2020) 106466

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: http://www.elsevier.com/locate/comphumbeh

Full length article

Communicative and social consequences of interactions with voice assistants

Sara Aeschlimann [a], Marco Bleiker [a], Michael Wechner [b,c], Anja Gampe [a,*]

[a] Department of Psychology, University of Zurich, Binzmuehlestrasse 14, 8050, Zurich, Switzerland
[b] Wyona, Stockerstrasse 32, 8002, Zurich, Switzerland
[c] Netcetera, Zypressenstrasse 71, 8040, Zurich, Switzerland

ARTICLE INFO

Keywords:
Voice assistants
Artificial intelligence
Communication
Social
Cooperation

ABSTRACT

The growing prevalence of artificial intelligence and digital media in children's lives provides them with the opportunity to interact with novel non-human agents such as robots and voice assistants. Previous studies show that children eagerly adopt and interact with these technologies, but we have only limited evidence of children's distinction between artificial intelligence and humans. In this study, the communication patterns and prosocial outcomes of interactions with voice assistants were investigated. Children between 5 and 6 years ($N = 72$) of age solved a treasure hunt in either a human or voice assistant condition. During the treasure hunt, the interaction partner supplied information either about their knowledge of or experience with the objects. Afterwards, children were administered a sharing task and a helping task. Results revealed that children provided voice assistants with less information than humans and that only the type of information given by a human interaction partner was related to children's information selection. Sharing was influenced by an interaction between type of information and interaction partner, showing that the type of information shared influenced children's decisions more when interacting with a human, but less when interacting with a voice assistant. Children in all conditions enjoyed the treasure hunt with the interaction partner. Overall, these results suggest that children do not impose the same expectations on voice assistants as they do on humans. Consequently, cooperation between humans and cooperation between humans and computers differ.

1. Introduction

We interact with humans, digital devices and artificial intelligence devices in our everyday lives in the digital society. Consequently, children are exposed to digital media and digital devices from an early age. The growing prevalence of artificial intelligence and digital media in children's lives provides them with the opportunity to interact with novel non-human agents such as robots and voice assistants. Educational robots, for example, support children in learning by serving as a tutor (Kennedy et al., 2015) or a teaching assistant (Chang et al., 2010). Furthermore, robots assist children with physical disabilities to move and play, and they help children with autism learn about communication and emotions (Liu et al., 2008). Previous studies show that children eagerly adopt and interact with these technologies (Lovato & Piper, 2015; McReynolds et al., 2017). While we as adults have a history of interactions with humans and artificial intelligence devices and can tell the differences between them, what about children growing up in the digital age? Does it matter with whom they interact? Do they impose the same expectations on humans as they do on virtual agents? How do children distinguish between artificial intelligence and humans?

In this study, we report on the effects of children's interactions with voice assistants. Voice assistants are speech-based user interfaces that interpret human language and offer a number of services, such as playing music and writing lists or messages (Hoy, 2018). Already one in four children between 5 and 16 years of age live in a household with a voice-activated virtual assistant in the UK (CHILDWISE Monitor, 2019). However, little is known about how children perceive these devices. Are they perceived similarly to humans as communicative and social interaction partners despite their features being very reduced?

* Corresponding author. University of Zurich, Department of Psychology, Binzmühlestr. 14, Box 21, 8050, Zurich, Switzerland.
E-mail address: a.gampe@psychologie.uzh.ch (A. Gampe).
URL: https://osf.io/4r3c8/?view_only=0e514d8a4fb543c8a1370af8bf4976eb (A. Gampe).

https://doi.org/10.1016/j.chb.2020.106466
Received 7 January 2020; Received in revised form 13 June 2020; Accepted 21 June 2020
Available online 27 June 2020
0747-5632/

2. Theoretical perspective: cooperation in communication and interactions

Our human society is built on the principle of cooperation (Fehr & Fischbacher, 2004; Kohn, 1992; Lindenfors, 2017). We cooperate in social groups, with family and friends, as well as with people we have not met before. We are cooperative in our communications and social interactions based on a set of widely shared beliefs that prescribe how individual group members ought to behave in a given situation (Fehr & Fischbacher, 2004). Cooperation is seen as a mutualistic collaboration since it enables individuals to put their heads together to create complex solutions (Tomasello et al., 2012). The cooperative structures of communication and social interactions are said to be closely related: Human cooperative communication comprises both social-cognitive skills for creating the joint intention and attention needed in verbal communication and the prosocial motivations needed in social interactions to help and share with others (Tomasello, 2008). Mutuality and reciprocity are thus core principles of our interactions with others.

But how do we cooperate with interaction partners that are not human, such as artificial intelligence devices? Building on earlier work by Moreno and Mayer (2000) on social cues in computer-based environments and work by Reeves and Nass (1996) on how people interpret computers as social partners, Mayer, Sobko and Mautone proposed *social agency theory* (2003). Social agency theory essentially argues that the social cues of a computer (e.g., modulated intonation, human-like appearance) encourage people to interpret the interaction with a computer as being social in nature. Once the interaction is interpreted as social interaction, the social rules of human-to-human communication come into play, leading people to act as if they were in a conversation with another human person. Thus, people are more likely to engage in deep cognitive processing to make sense of what the artificial agent is saying and communicate accordingly.

Given these assumptions, voice assistants should be viewed as interaction partners with reduced human features. Although voice assistants are able to interpret human language and can reply using synthesised voices, voice assistants lack all kinds of non-verbal communicative aspects such as gaze, facial expressions, postures, gestures and emotional reactions. For this reason, the communicative means of voice assistants are reduced compared to those of humans. Moreover, they can only communicate with synthesised voices, which leads to reduced variability in their intonation. For all of these reasons, communication with voice assistants is more challenging, and the cooperation in communication between humans and voice assistants would likely be limited. Voice assistants do not have physical appearance and presence, and they are not able to move or act. Voice assistants, therefore, have a limited repertoire for real-world interactions. In terms of behavioural interactions, reactions to user prompts cannot be contingent on the behaviour of the user. Cooperation in social interactions, i.e., in terms of achieving a joint goal, is, therefore, also limited in voice assistants. Taken together, these features increase the social-cognitive skills humans need when attempting to create joint attention and intention in interactions (Tomasello, 2008) with voice assistants. It is expected it would be harder to perceive voice assistants as social actors than, i.e., social robots.

Given the modified communication styles of voice assistants on the one hand, and their altered behaviours, on the other hand, it is especially likely that communication with voice assistants differs from that with human interaction partners. Expanding our knowledge about changes in communication styles in children who interact with voice assistants, as well as the related social outcomes, is of utmost importance in the current information society (Webster, 2002). In this study, we thus aim to answer the following questions:

RQ1 What do children communicate with a voice assistant?

The content of children's communication with a voice assistant is the first step in investigating the cooperation principle in communication. If children differentiate what kind of information they provide to a human or a voice assistant, we can infer that two different principles of cooperation exist.

RQ2 How does the interaction with a voice assistant affect children's subsequent social interactions with humans?

Due to the above-mentioned limited possibilities to cooperate with a voice assistant in a social interaction, we aim at shedding light on the cooperation by investigating the consequences for social interactions with humans. This is important to investigate if children are exposed to and interact with voice assistants often. If children share and help differently after interactions with voice assistants, we can infer that cooperation in the two types of interaction differs.

3. Related literature: communicative and social interactions with voice assistants

In the domain of communication, advances in natural language processing allow voice assistants to process user requests remarkably fast. Nevertheless, pragmatic and discourse skills are still limited (Hirschberg & Manning, 2015). Adults adjust to these limitations when communicating with voice assistants, e.g., by shortening sentences, simplifying their language, increasing the volume and using repetition (Pelikan & Broth, 2016). In contrast, while children can adapt their mode of communication with voice assistants (Beneteau et al., 2019), they still require help from adults in formulating adequate questions to ask them (Yarosh et al., 2018). When integrated into family life, communication breakdowns with voice assistants occur that require repair strategies and discourse scaffolding by other family members (Beneteau et al., 2019). In fact, children perceive voice assistants as a less appealing and pleasurable interaction partner (Sinoo et al., 2018). Some social cues seem to play an important role in the communicative expectations of children in interactions with robots: children are particularly attentive and receptive to robots with high non-verbal contingency (Breazeal et al., 2016) and an expressive narrative style (Westlund et al., 2017) and show longer engagement when robots adapt to their affective states (Ahmad et al., 2017).

Given these previous research designs, we cannot disentangle whether children assume voice assistants to be human-like communication partners and do not want to adapt their communication style. Nor can we determine whether children's social-cognitive adaptation skills to artificial intelligence remain tenuous, and the contingent cues provided by voice assistants are insufficient for them. Furthermore, we do not know whether having more human-like features in interactions would increase communication with a voice assistant.

In the domain of social interaction, a growing body of literature focuses on how people conceptualise artificial agents in terms of interaction partners. One of the most striking findings is that people treat computers as social actors (Isbister & Nass, 2000; Nass & Moon, 2000; Reeves & Nass, 1996; Sproull et al., 1996). Research with (social) robots provides evidence for children's conceptualisation of artificial agents. A majority of children believe that robots are social beings (e.g., someone who can be a friend, play, comfort and keep secrets) with mental states (e.g., think, be intelligent, be happy), emotional states (e.g., be sad, like) and, to some extent, even moral states (Beran et al., 2011; Kahn et al., 2012) – even if the robot is not a humanoid (Melson et al., 2009). Children also behave socially towards robots (e.g., hugging, correcting, shaking hands, making eye contact; Kahn et al., 2012), react with stress when robots are mistreated or behave incorrectly (Kahn et al., 2006; Melson et al., 2009; Kahn et al., 2012) and give in to pressure to socially conform with a group of robots (Vollmer et al., 2018). Even small children quickly form social bonds with them – just as they form bonds with other people (Tanaka et al., 2007). These results support the social agency theory.

However, children also show some form of differentiation between human and artificial agents. Children engage in more physical and visual contact with living agents than social robots (Melson et al., 2009; Shahid et al., 2014). They feel a greater sense of friendship with living agents than they do with virtual agents (Sinoo et al., 2018) and associate more morality and mental states with living agents (Melson et al., 2009). Furthermore, children aged 4–10 years afforded robots less moral concern than living agents (Sommer et al., 2019). Beyond these results, studies with adults showed that cooperation with a human occurred more frequently than with a robot, and adults showed more fairness towards humans, although the rates of reciprocity were similar with both kinds of agents (Sandoval et al., 2016).

Given the previous findings, we have limited understanding of the cooperation between children and voice assistants. Although we have some insights into conceptualisations of social robots by children, we cannot transfer the results to voice assistants due to their limited human-like features. We thus need a study examining cooperation with voice assistants in social interactions.

To investigate the open research questions, what children communicate with a voice assistant and whether the interaction affects children's subsequent social interactions with humans, we developed a new paradigm to reveal the principles of cooperation between children and voice assistants. Children were administered a treasure hunt game with either a human or a voice assistant (interaction partner) that provided knowledge-based or experience-based information about the treasure hunt cues (information type). In the second part of the study, children were administered tasks on sharing (Fehr et al., 2008; Moore, 2009) and helping (Beier et al., 2019; Warneken & Tomasello, 2006) to investigate the prosocial consequences.

H1 Children communicate with the voice assistant providing experience-based information similarly than with humans, more so than they do with the voice assistant providing knowledge-based information.

Talking about personal experiences hints at the capacity for joint attention and prosocial motivations in interactions with voice assistants with increased human-like features. If children expect this to be a pre-requisite of cooperative communication, communication should increase with virtual agents providing experience-based information compared to those providing knowledge-based information and be more comparable to a human interaction partner.

H2 Children share and help less after interactions with voice assistants compared to humans.

In terms of social consequences, it is likely that children differentiate between the voice assistant and the human interaction partner given previous findings in adults (Sandoval et al., 2016) and children interacting with social robots (Melson et al., 2009). Having more human-like features, as proposed by the social agency theory, would result in higher prosocial rates in the experience condition.

4. Method of current study

To investigate the research questions, children were administered two tasks, a treasure hunt task and a prosocial task. In the treasure hunt task, children interacted with two human experimenters in one of four conditions. The conditions were a combination of the **interaction partner** (a human or a voice assistant) and the **information type** (knowledge or experience) during the treasure hunt, resulting in four between-subjects conditions (human – knowledge, human – experience, voice assistant – knowledge, voice assistant – experience). While human experimenter 1 (E1) interacted with the child in the study room during the familiarisation for the treasure hunt, human experimenter 2 (E2) was only present via her voice in all conditions. In the voice assistant condition, the interaction partner E2 was introduced as a voice assistant, while in the human condition, the interaction partner E2 was introduced as a friend of E1 (interaction partner). After the introduction, children completed a treasure hunt together with E2. During the treasure hunt, E2 either provided personal experience-based or knowledge-based information about the treasure hunt objects (information type). In the second part of the study, children were administered tasks on sharing (Fehr et al., 2008; Moore, 2009) and helping (Beier et al., 2019; Warneken & Tomasello, 2006) to investigate the prosocial consequences.

4.1. Participants

We included 72 5- to 6-year-old children, 18 in each of four conditions. The final sample was a matched sample of $n = 72$ taken from a whole sample of 83 tested children. This sample was extracted by creating similar background environments for the children in the four conditions based on important covariates: age of the children at test time (we included children between 5 and 6 years of age, which is very broad), parental education (parental education is associated with children's cognition), media use of the child and the parents (the behaviour of the children in the treasure hunt task depends on the experiences of the children with digital media in general and smart home gadgets in particular), the social dimension of the Strengths and Difficulties Questionnaire (Goodman, 1997; Klasen et al., 2003) and a baseline need-of-help recognition task (Brielmann & Stolarova, 2014) (these two latter scores were included to have a baseline of social behaviour independent of the manipulation during the treasure hunt, see how they were assessed below). Refer to Table 1 for an overview of the demographics.

Parents filled in a media consumption questionnaire that assessed parent and child media frequency (Common Sense Media, 2013; Genner et al., 2017). Parents were asked how often they use computers, smartphones, tablets, TV and voice assistants on a 7-level scale (never, once a month, once a week, several times a week, once a day, several times a day, hourly) and how many apps they have installed on their smartphone (none, 1–4, 5–9, 10–19, 20–29, $\geq$30 apps). The levels of these questions were summed up as a parent media score. Parents were asked how often their child uses computers, smartphones, tablets, TV and voice assistants on the same 7-level scale (never, once a month, once a week, several times a week, once a day, several times a day, hourly) and to report the time in minutes that children spent with those devices yesterday. Furthermore, parents were asked how frequently children use the electronic devices for specific activities (play games, listen to music or stories, watch videos on the internet, receive/send messages, edit pictures/videos, phone, search for information on the internet, send photos/videos, listen to the radio, watch TV, use a voice assistant) on a 5-level scale (never, seldom, sometimes, often, very often). The levels of these questions were summed up as a *child media score*.

We used a modified version of the Need-of-Help-Recognition Task (Brielmann & Stolarova, 2014). Children sorted cards showing everyday situations of someone either achieving a goal or needing help achieving a goal onto two stacks. The task consisted of 14 trials; two practice trials (demonstrating the procedure of the task), followed by 12 test trials, in which the child could sort a card onto one of two stacks: the person needs help, or the person does not need help. We coded for each test trial whether children sorted the card correctly (1; the person in the picture does not need help and child sorted the card onto the does not need help stack; the person in the picture does need help and child sorted the card onto the does need help stack) or not (0). The number of correct sorting choices was summed up as the need-of-help recognition score.

All procedures were approved by the ethics committee of the University of Zurich and performed in accordance with the ethical standards of the 1964 Helsinki declaration and its later amendments. All parents gave informed consent. Children received a small toy and a certificate for their participation.

4.2. Treasure hunt task

Familiarisation. For the treasure hunt, 19 coloured treasure boxes were hidden in a room under a table, in shelves and behind chairs: Eight boxes were target boxes, while the other 11 boxes were used as distractors. Two of the eight target treasure boxes were used to explain the procedure of the treasure hunt to the child: Children started with a key

Table 1
Demographics of participants.

	Human - Experience	Human - Knowledge	Voice Assistant - Experience	Voice Assistant - Knowledge	Group Comparison
Girls/Boys	7/11	10/8	10/8	12/6	$X^2 = .89$, $p = .35$
Age (Days)	2065.39	2042.94	2051.61	2039.17	all $p \geq .35$
Education Father	4.56	3.83	4.25	4.18	all $p \geq .13$
Education Mother	4.11	4.0	4.22	4.35	all $p \geq .53$
Child Media	1.78	1.83	1.61	1.94	all $p \geq .39$
Parent Media	4	4.28	3.78	3.78	all $p \geq .45$
SDQ: Social	7.28	7.06	7.44	7.5	all $p \geq .63$
Need-of-Help Recognition Task	9.78	10.22	10.5	9.61	all $p \geq .11$

that opens a box, and the experimenter knew where the treasure box was that had to be found next. Once children found the box, there was a cash box underneath the treasure box. This cash box could be opened with the key the child was given, and inside it, the next key was hidden. After two trials, E1 pretended not to know which cash box the second key would open, but they mentioned that E2 had done the treasure hunt several times before.

Introduction of E2. E1 then introduced E2, called "Sila", to the children by giving them information about their new interaction partner in equal domains but with different content, see Table 2. E1 subsequently presented the child with a small black box with which they could connect to Sila. Sila was talking as an audible voice via loudspeakers. Children were encouraged to ask a few specific experimenter-prompted questions to Sila that were answered differently by both interaction partners to further help children understand the nature of the interaction partner, see Table 2. If children were too shy to ask, E1 asked the questions. Sila as a voice assistant spoke with a monotone voice, while Sila as a human spoke with a modulated human intonation. We measured how many questions the child asked Sila as a measure of willingness to communicate with the unfamiliar interaction partner (*questions*).

Table 2
Introduction of the interaction partner (Sila).

Phase	Question	Voice Assistant	Human
Introduction of E2		"Sila is not a man, she was assembled by a man."	"Sila is a good friend of mine."
		"She never has to sleep or eat, she just needs electricity."	"She likes to sleep in the morning and eat a lot for breakfast."
		"She can't move, she can only talk to you."	"She likes to cycle and hike outside."
		"She always does what you tell her and can help with many things."	"She works as a cook and loves to prepare spaghetti."
		"for example, she can tell you when the next train's leaving, or tell you where to go when you're driving."	"She loves to ride the train but sometimes she also loves to travel by car."
		"She can also play music…	"She likes to listen to children's songs."
		… or say what the weather will be like tomorrow."	"She doesn't like bad weather at all."
Question phase	"What's the weather like with you?"	"There's a 95% chance it's cloudy and a 90% chance there'll be 10 mm of rain in the next 2 h."	"I don't know, there are dark clouds, so it might rain soon, but I didn't look at the weather forecast."
	"What's your favorite color?"	"My favorite color is a shade of green but with several dimensions."	"My favorite color is light green."
	"What's your mother's name?"	"I don't have a mum."	"My mum's name is Daniela."
	"How old are you?	"I've been on for 3445 h."	"I will turn 30 years this year."
	"Do you have time for the treasure hunt?"	"I am always at your service."	"I have half an hour left, then I have to go on."

Interaction with E2. After this familiarisation, E1 brought the child to the treasure hunt part of the room and explained that the child only needs to tell E2 what is depicted on the key. Then, E1 left that part of the room. Each time the child mentioned the picture on the key, E2 provided two sentences with information on the object of the key and told the child which box to look for, see Fig. 1. The sentences were chosen in a way that the experience-based information contained their own experiences of senses and emotions, while the knowledge-based information contained objective facts. These differences between information types were the most cited differences between humans and voice assistants in an adult survey of $n = 68$ participants that was administered prior to the study. They testify to the importance of fostering joint attention and intention as well as prosocial motivations in interactions with others in human adults (Tomasello, 2008) and increase the number of human-like features.

In the treasure hunt task, we measured how often children talked to the interaction partner during the search phases in which information sharing was optional but not obligatory: children talked mainly about the search process ("I can't find the box") and about the opening of the boxes ("Should I open the box?") or they did not communicate with E2 (*communication*), $\kappa = 0.91$. Verbal attempts/reciprocity was used as a measure for communication (Melson et al., 2009).

Interview. After the treasure hunt was solved, E1 came back and asked the child questions about E2 to check whether the child realised which interaction partner they interacted with during the treasure hunt. The questions E1 asked were: "Is Sila a human or built by a human?", "Has Sila a mother?", "Can Sila be sad?", "Can Sila be scared?", "Did you like the treasure hunt with Sila?", "Would you like to play again with Sila?", "Would you share your favorite toy with Sila?". For the interview questions, we coded an affirmation (yes) or negation (no) as well as undecided (don't know) for each question (*interview*), $\kappa = 0.96$.

4.3. (Pro-)Social tasks

Sharing. We used a Resource Allocation Task (Fehr et al., 2008; Moore, 2009; Williams et al., 2014) with four trial types. The pictures for this task were chosen from Egger et al. (2011), and the task consisted of 18 trials in which children were asked to choose between two different fixed options to distribute stickers between themselves and other kids. Two of these trials were practice trials, and in the first of which, children could either choose one sticker for themselves and three for the other children or two stickers for both. In the second practice trial, children could either choose three stickers for themselves and none for the others or three stickers for both (demonstrating the format of the task). The practice trials were followed by two repetitions of each trial type, which offered the child a forced choice between two alternative distributions of stickers. In no cost sharing trials, children chose between the allocation (1, 1) and (1, 0) – either one sticker for themselves and one for the other child or one for themselves and none for the other child. In costly sharing trials, children chose between (1, 1) and (2, 0), in no cost envy trials,

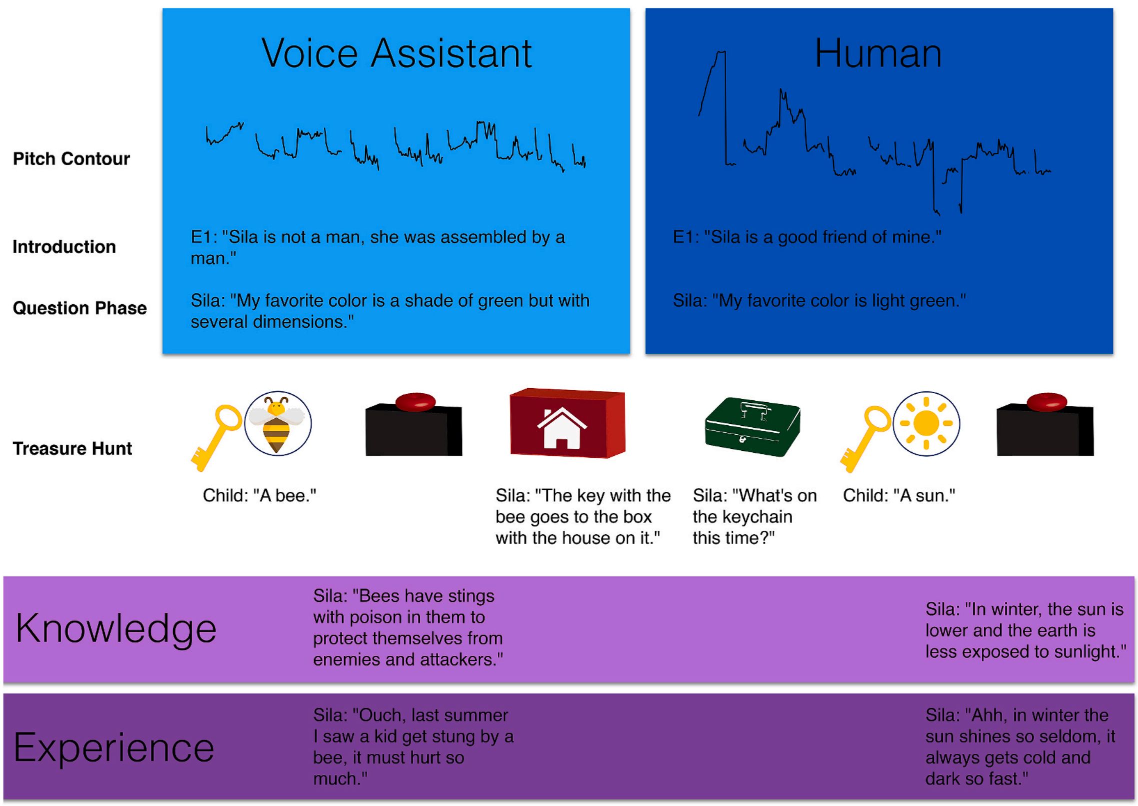

Fig. 1. Design and procedure of the study. Children were tested in one of four between-subject conditions interacting with either a voice assistant or a human that provided either knowledge-based or experience-based information during the treasure hunt. Differences between interaction partners were operationalised by monotone vs. dynamic intonation, as can be seen in the pitch contour and the different sentences spoken during the introduction and question phase. Differences in information types were operationalised by the different sentences about the objects on the keys during the treasure hunt.

between (1, 1) and (1, 2), and in costly envy trials, between (1, 1) and (2, 3). We coded for each trial whether children chose the prosocial option (choosing the (1, 1) distribution in the no cost sharing and the costly sharing, or the (1, 2) option in the no cost envy, or the (2, 3) option in the costly envy trials) or not (*sharing*), κ = 0.96.

Helping. Three helping trials were administered while the child was watching an episode of "Shaun the Sheep" to increase the costs for helping (Warneken et al., 2007) as preschoolers are nearly at the ceiling level in helping tasks at this age. In three trials, E1 dropped 10 pens, 10 office clips and 10 sheets of paper and reached for them from a chair but could not pick them up (Beier et al., 2019). After the occurrence of the problem in each trial, there were three phases: E1 focused on the object and made three unsuccessful attempts to reach the objects, saying "Ohh!" each time (1–10 s), then alternated her gaze between object and child (11–20 s). In addition, she verbalised her problem while continuing to alternate her gaze (21–30 s).

We used the same coding scheme from 1 to 6 points for each trial in line with Beier et al. (2019) in their study: Scores reflect the speed and the presence of helping behaviour (range = 1–6). Children who helped received scores of 5 or greater, children who did not physically help but offered verbal solutions or other acknowledgments of the problem received scores from 2 to 4, and children who neither helped nor commented supportively received a score of 1 *(helping)*, κ = 0.92.

5. Results

5.1. Communication with the interaction partner

During the introduction of E2 as an interaction partner for the treasure hunt, we measured how many questions children asked E2 (about their favorite color, their mother's name and their age). Children asked the same number of *questions* to the voice assistant (1 question $n = 4$; 2 questions $n = 1$; 3 questions $n = 18$) and the human (1 question $n = 3$; 2 questions $n = 3$; 3 questions $n = 14$), X^2 $(2, N = 72) = 1.44, p = .49$.

During the treasure hunt, we were interested in what information children would exchange with the interaction partner. All children told the interaction partner in all trials about the object on the key. For the communication during the search phases, we investigated whether children exchanged information with the interaction partner differently in the four conditions. We ran a logistic regression with the interaction partner, information and trial as fixed effects, child media as the covariate and the subject as a random effect. We then predicted whether children would say anything to the interaction partner (no communication, communication). We found that children were less likely to communicate with the voice assistant, and they were less likely to communicate when the interaction partner provided experience-based information compared to knowledge-based information (Fig. 2). Furthermore, we observed significant interactions between the interaction partner and information, suggesting that the difference between the information types was greater when interacting with a human, see Table 3.

5.2. Interview about the interaction partner

After the treasure hunt, we asked children about their interaction partners to ensure that they knew with whom they were interacting and determine how much they liked the treasure hunt. We found that children gave different answers to the questions (human, mother, sad, scared), showing that they knew when they interacted with a human vs. a voice assistant. All children across all conditions equally enjoyed the treasure hunt; see Table A2 and Fig. 3.

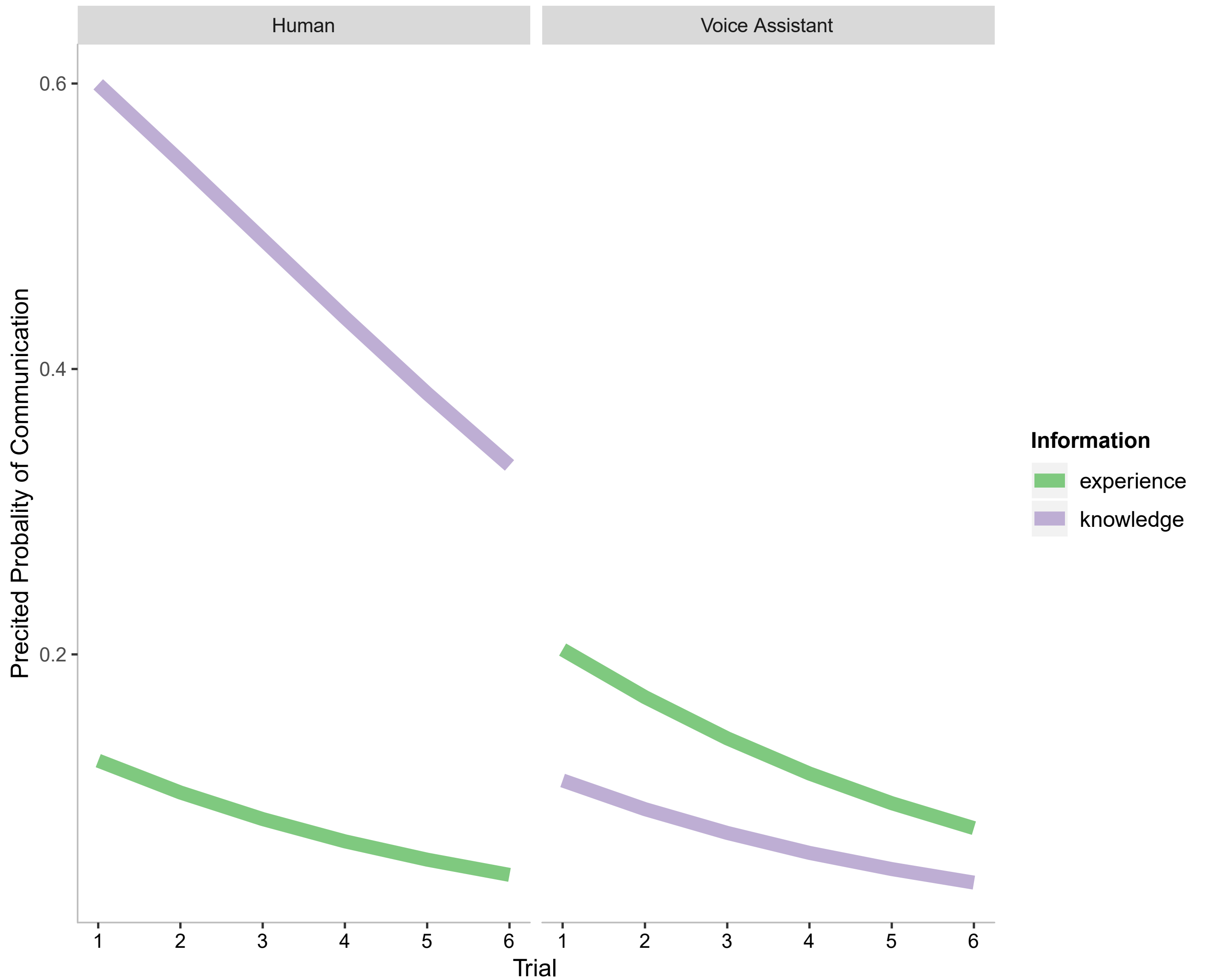

Fig. 2. Predicted probabilities of children communicating with the interaction partner during the search.

5.3. (Pro-)Social tasks

After the treasure hunt, we were interested in how children would decide on sticker distributions for themselves and absent, same-aged interaction partners that were presented in pictures. We investigated whether prosocial choices varied in the four conditions. We ran a logistic regression with the interaction partner and information as fixed effects, child media as a covariate and the subject and condition as random effects. We then predicted whether children chose the prosocial option (yes, no). We found that children were less likely to choose the prosocial option when the interaction partner provided knowledge-based information. Furthermore, we found significant interactions between the interaction partner and information, suggesting that the difference between the information types was greater after interacting with a human, see Table 3 and Fig. 4.

In the helping task, we were interested in whether children would change their prosocial behaviour towards the experimenter. Seven children were excluded from this analysis due to video loss ($n = 2$) or parental/sibling interference ($n = 5$). We ran an ordinal regression with the interaction partner and information as fixed effects, child media as the covariate and the subject and condition as random effects. We then predicted how children reacted in this task. None of the predictors was significant in explaining variance in the helping task.

6. Discussion

This study investigated children's communicative interactions and the social consequences of interactions with voice assistants. Using a combination of information (knowledge, experience) and interaction partners (human, voice assistant) as characteristics, we had children perform a treasure hunt firstly and two social tasks subsequently. Results show that children talked less to the voice assistant than to the human during the treasure hunt. After children interacted with partners providing personal experiences, they said fewer sentences but shared more prosocially. During both the treasure hunt and the sticker task, interaction effects between the interaction partner and information were significant, showing that the type of information shared influenced children's decisions more when interacting with a human, but less when interacting with a voice assistant. We found there were no predictors of children's helping behaviour.

6.1. Communication

The results of the current study provide two novel contributions to

Table 3
Results of the logistic regressions on communication and sharing and the ordinal regression on helping.

Domain of Test	Predictor	Estimate	*SE*	*p*-value
Communication	Intercept	−0.32	0.85	0.71
	Interaction Partner (voice assistant)	−2.48	1.05	**0.02**
	Information (experience)	−2.34	1.08	**0.03**
	Trial	−0.22	0.09	**0.02**
	Child Media Score	0.52	0.22	**0.02**
	Interaction Partner x Information	3.05	1.53	**0.05**
Sharing	Intercept	−0.26	0.61	0.67
	Interaction Partner (voice assistant)	0.46	0.42	0.27
	Information (experience)	0.97	0.43	**0.02**
	Child Media Score	0.10	0.09	0.27
	Interaction Partner x Information	−1.18	0.60	**0.05**
Helping	Interaction Partner (voice assistant)	−0.82	1.84	0.65
	Information (experience)	−2.29	1.97	0.25
	Child Media Score	0.38	0.39	0.33
	Interaction Partner x Information	0.58	2.62	0.83

our understanding of children's communicative skills. They show that children share less information with voice assistants than they do with humans and that the type of information given by a human interaction partner is related to children's information selection. Although all children provided the required and essential information about the objects on the keys, children selectively gave some partners more information about their search process. Results of the first trial revealed that with not even a 20% probability, children told a voice assistant about their search progress, but told a human with a 60% probability. Previous research showed that children need scaffolding during challenging communications with voice assistants (Beneteau et al., 2019; Yarosh et al., 2018), leaving two possible explanations open: either children do not know how to appropriately talk to a voice assistant, or they do not differentiate between voice assistants and humans in their communicative expectations. The results of the current study provide evidence that children make a clear distinction between humans and voice assistants as they assume voice assistants do not need to be informed about the same kinds of things as humans. They show a third option: Children know that voice assistants do not care much about progress talk while humans like to keep informed. Therefore they know how to talk with voice assistant appropriately. This option is plausible, because we do know that they are able to provide more information, as the children in our study did so in the human condition. These results might imply that the cooperative structure of communication is more difficult to establish for voice assistants due to the missing features enabling joint attention and joint intention (Tomasello, 2008). However, since the human in this study possessed the same features, it is likely that the cooperative structure is different for voice assistants than for humans. This leads to a decrease in the number of communicative acts directed at artificial intelligence.

The study adds to our knowledge of pragmatic development and shows that children adapt their communication style according to their interaction partner. Cooperation in communication requires us to be as informative as needed, to be truthful and clear and to be relevant (Grice, 1989), known as the Gricean maxims. The information the interaction partners gave during the treasure hunt was in all conditions irrelevant and violated the quantity maxim (do not provide more information than is required). For this maxim to be fulfilled, the interaction partner only needs to tell the children which box to look for next without providing information about the object on the key. Beyond violating the quantity maxim, the experience condition was poorer in the maxims of relevance and quality. Sharing their own experiences is less relevant than talking about the appearance and function of the objects. We have evidence that children at preschool age already recognise and penalise violations of the Gricean maxims (Davies & Katsos, 2010; Eskritt et al., 2008; Gweon et al., 2014). Expecting relevant and quantitatively well-suited information, children in our study did share more information about the treasure hunt when confronted with a partner that shared more relevant information (knowledge information). But children did distinguish less between the two types of information in the voice assistant condition, providing further evidence that children have different pragmatic expectations of artificial intelligence. Increasing the information about joint attention experiences and prosocial motivations did affect children's communication with the human interaction partner but did not affect their communication with the voice assistant. These findings contrast to the assumptions of the social agency theory that proposes stronger human-like behaviour if more human-like features are given to an artificial agent. It might, however, be that a certain threshold of human-like features must be met to communicate with artificial agents similarly to human agents. If this is the case, then voice assistants might fall below this threshold because they do not have a sufficient number of communicative cues and lack all kinds of non-verbal communicative aspects such as gaze, facial expressions, postures, gestures, emotional reactions and behavioural actions. The results provide further evidence that the cooperative structure in communication differs between

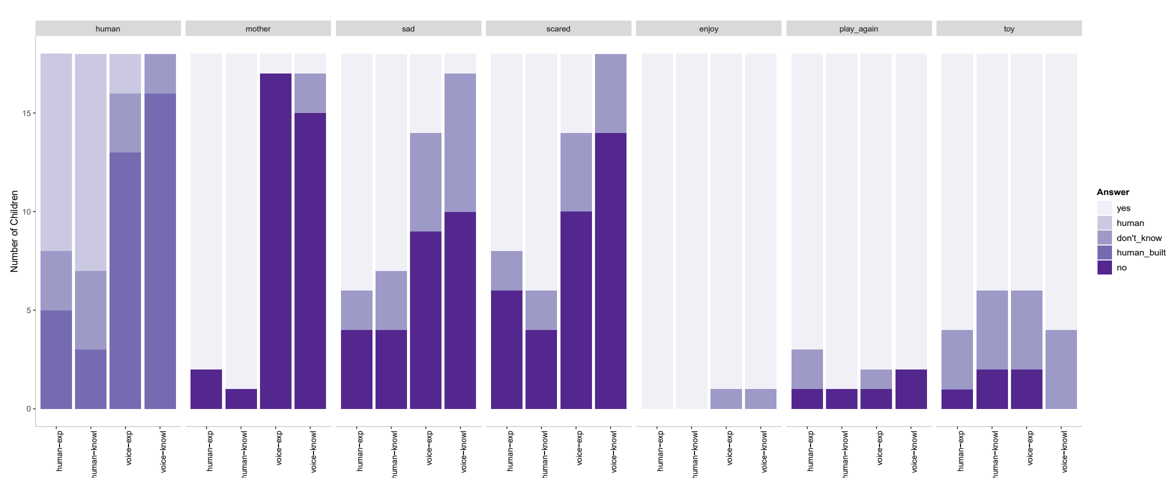

Fig. 3. Answers children gave to the interview questions. Children remembered which interaction partner they did the treasure hunt with, giving varied answers across the interaction partners but equally enjoying the treasure hunt.

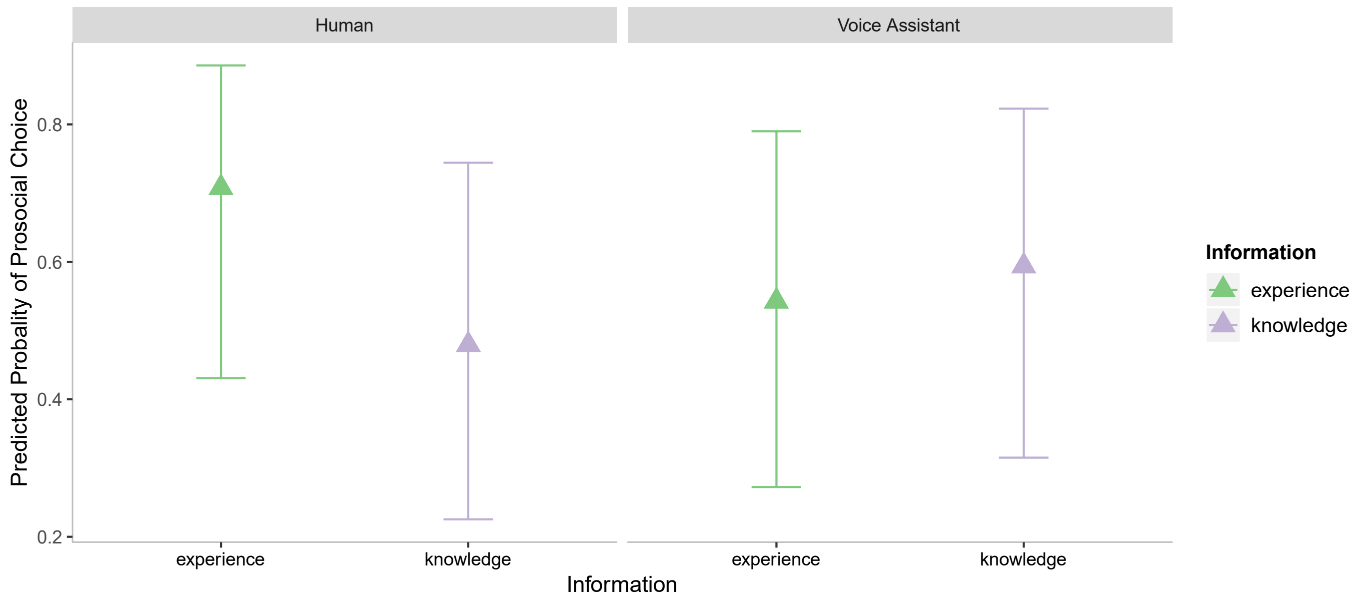

Fig. 4. Children's predicted probabilities for prosocial choices during the sticker task.

humans and voice assistants.

6.2. Social consequences

In contrast to the findings in the domain of communication, sharing experience information during the treasure hunt influenced children positively in the prosocial distribution of stickers. This finding converges with earlier findings and theoretical accounts that emotional arousal and relatedness are linked to prosocial decisions in adults and children: empathic concern for the interaction partner was a major predictor of greater sharing in the sticker task already at three years of age (Bastian et al., 2014; Decety et al., 2016; Pavey et al., 2011; Williams et al., 2014). A novel contribution is that shared auditory information about previous experiences is related to a higher probability of prosocial sharing decisions. This pattern was, however, not shown in the voice assistant condition, again suggesting that children differentiate between humans and voice assistants in their social interactions. The reduced scope to really interact with a voice assistant might lead to decreased prosocial motivations. This, in turn, may result in more selfish decisions among children.

The results from the question of how many children enjoyed the treasure hunt with Sila provide further evidence that children perceive artificial intelligence agents as social beings (e.g., someone who can be a friend, gives comfort and keeps secrets) and attribute mental states (e.g., someone who is intelligent and has feelings) and, to some extent, even moral standing to them (e.g., someone who deserved fair treatment and should not be harmed) (Beran et al., 2011; Kahn et al., 2012; Melson et al., 2009). Evidence for social interactions with robots shows that children quickly form social bonds with them (Tanaka et al., 2007) and treat them socially (e.g., hugging, handshaking, joint attention, prosocial behaviour, social conformity) (Kahn et al., 2012; Kim et al., 2009; Melson et al., 2009; Vollmer et al., 2018). Children are particularly engaged when interacting with a robot with high non-verbal contingency (Breazeal et al., 2016), an expressive narrative style (Westlund et al., 2017) and the ability to adapt to their affective states (Ahmad et al., 2017). The social agency theory (Mayer et al., 2003) argues that the quality of social interaction increases with the number of social characteristics an artificial agent possesses because humans tend to interpret the interaction as a social communicative situation, which leads to a deeper processing of the information presented by the agent. In particular, when interacting with artificial agents with more social characteristics, children have difficulties in correctly classifying them as artefacts (Kahn et al., 2013; Turkle, 2005). However, since research to date has only been conducted with embodied artificial agents, future research has to investigate the perceptions of disembodied agents such as voice assistants, as well as the consequences for cooperation.

7. Implications

Every invention has the potential for improvement. For example, cars became safer, and mobile phones became more multifunctional. Voice assistants can be a great tool to support humans in their daily activities. In particular, they have vast potential in supporting children's learning and development. Children learn by imitation, and, therefore, we encourage product designers and developers of voice assistants to consider the implications of our findings. Improving voice assistants can help maximising the favorable outcomes of interactions between children and voice assistants. Our study shows that children do hold different conceptualisations of voice assistants and humans and that adding human features (e.g., sharing experience-based information) does not affect children's communicative or social behaviour in the short run. These insights are helpful for the development of future voice assistants because they reveal that product designers do not need to invest in these features. However, the reduced number of instances in which children shared information about their environment during the search in all conditions is alarming. If we imagine children to have regular interactions with voice assistants starting from an early age, we could conceive a world in which children no longer talk much about the things they encounter. They would be used to talking to voice assistants and may assume that humans behave similarly. This would be a very silent world. These possibilities support the relevance of either reducing children's early contact with voice assistants or improving the communicative interaction between voice assistants and children substantially.

8. Limitations and future research

We did not find any differences between the conditions for the helping task. This task showed only small variability, as only a few children diverted their attention from the video. Therefore, further research is needed to investigate the consequences of interactions with voice assistants with tasks that provide a more detailed picture of prosocial willingness. Also, perhaps, more prolonged interactions with voice assistants can be designed to see if the influence does not exist when helping or not after such a short manipulation in which all children finally receive the reward in the treasure hunt. Having different outcomes for the treasure hunt could be a further step.

An additional venue for future research is to examine the possible effects of different brands. We used a non-brand neutral device that was novel to all children to control for varying previous experiences. Our study was performed by using a human imitating a voice assistant. We did not use any particular existing voice assistant devices, such as, for example, Amazon's Alexa or Apple's Siri, because the development of an appropriate and controllable voice assistant capability was outside the scope of our study. It would be interesting, though, to categorise the various existing voice assistants and compare their status and behaviour in a future study. It is possible that children who are already familiar with one of the existing voice assistants would communicate differently with this device than to another device and/or with our non-brand assistant. To investigate this line of research, studies manipulating the familiarity and visual appearance of the devices are needed.

Furthermore, only the short-term effects of a single interaction with voice assistants on prosocial behaviour were investigated experimentally. It would be interesting to investigate the effects of repeated interactions with voice assistants. Whether interaction with voice assistants is reflected in a child's cooperation skills in the long term will depend on various person- (e.g., age and gender of the child, context of use, supervision, extent of use, type of use) and device-specific factors (e.g., degree of possible interactivity, intended target group) (Biele et al., 2019). Previous research suggests that prosocial skills develop throughout childhood (Eisenberg & Miller, 1991), and prosocial behaviour increases (Eisenberg et al., 1991). Given this evidence, it is conceivable that the frequency of family and child media use may be related to the influence of the voice assistant. Especially in the case of younger children who grow up in an environment rich in media, repeated use of voice assistants could hinder their prosocial behaviour. In the future, it will be necessary to investigate more closely how forms of interaction differ in their long-term influence on cooperation and how specific media characteristics of children mediate this influence in different developmental phases.

Lastly, due to the complex design and procedure adopted in our study, we were not able to test a higher number of children in each condition. The robustness of the findings might, therefore, be limited and needs to be validated in future studies.

9. Conclusion

Overall, the current results and findings converge with previous findings on other artificial intelligence devices like humanoid and animaloid robots, showing that children like to play with such devices but do not impose the same expectations on them as they do on humans (Melson et al., 2009; Pelikan & Broth, 2016; Sandoval et al., 2016; Shahid et al., 2014; Sinoo et al., 2018; Sommer et al., 2019). These results provide significant insight into the influences of interactions with voice assistants on communication and prosocial behaviours. Children provided voice assistants with less information and shared stickers independent of the kind of information the voice assistant provided. These findings suggest the principles of human cooperation differs from the principles of human-computer cooperation.

Declaration of competing interest

None.

CRediT authorship contribution statement

Sara Aeschlimann: Conceptualization, Methodology, Formal analysis, Resources, Writing - original draft, Visualization. **Marco Bleiker:** Conceptualization, Methodology, Writing - review & editing, Project administration. **Michael Wechner:** Conceptualization, Methodology, Resources, Writing - review & editing. **Anja Gampe:** Conceptualization, Methodology, Formal analysis, Writing - original draft, Visualization, Supervision.

Acknowledgments

First and foremost, we thank the participants and their parents for their participation. We are also sincerely grateful to Sila Simmen and Ronja Barelli for their effort as second and third experimenter, respectively, playing the role of the interaction partner "Sila". We further thank Ronja Barelli for reliability coding. Last but not least, we would like to express our gratitude to Florence Bernays and Joanna Müller for their help with the graphics.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chb.2020.106466.

Appendix

Table A1
Information on the objects given by the interaction partner.

Object	Information	Experience
Bee	"Bees have stings with poison in them to protect themselves from enemies and attackers. They produce honey from pollen, which many animals eat."	"Ouch, last summer I saw a kid get stung by a bee, it must hurt so much. But honey buns are still very tasty."
Sun	"In winter, the sun is lower and the earth is less exposed to sunlight. Too much sun burns people's skin."	"Ahh, in winter the sun shines so seldom, it always gets cold and dark so fast. At least, you can't get a bad sunburn when you're outside."
Flower	"Flowers have bright colors to attract insects. For example, there are roses in different colors."	"Wow, the beautiful colors of flowers always put me in a good mood. Especially roses I find incredibly great".
Elephant	"Shortly after an elephant is born, it tries to get up and walk. Elephants have huge ears so that they can fan each other air."	"Aww, on my last visit to the zoo I saw a clumsy baby elephant stumbling after its mother. So sweet, his ears seemed far too big for the little body."
Plane	"Airplanes fly several hundred meters above the clouds. Thanks to the airplane, people can fly anywhere in a short time."	"What I particularly like about flying is that you can look down on the clouds. Two months ago I flew to America, which took a really long time."
Banana	"The riper a banana is, the softer it becomes. A banana contains sugar and other substances that supply the body with energy."	"Eww, there's nothing more disgusting than mushy bananas. But a lot of people like to eat them when they get tired."

Table A2
Results on the comparisons between answers in the interview questions. Interaction partner and information differences were analysed by Chi-Square tests. For both factors, Cochran-Mantel-Haenszel tests were run.

Question	Interaction partner Difference	Information Difference	Interaction partner and Information Difference
Is Sila a human or built by a human?	$X^2 = 29.43$, $p < .001$	$X^2 = 1.16$, $p = .762$	$X^2 = 28.71$, $p < .001$
Has Sila a mother?	$X^2 = 53.49$, $p < .001$	$X^2 = 2.29$, $p = .319$	$X^2 = 52.03$, $p < .001$
Can Sila be sad?	$X^2 = 18.94$, $p < .001$	$X^2 = 1.14$, $p = .567$	$X^2 = 18.65$, $p < .001$
Can Sila be scared?	$X^2 = 19.56$, $p < .001$	$X^2 = 0.27$, $p = .873$	$X^2 = 19.08$, $p < .001$
Did you like the treasure hunt with Sila?	$X^2 = 0.51$, $p = .473$	$X^2 = 0$, $p = 1$	$X^2 = 0.50$, $p = .480$
Would you like to play again with Sila?	$X^2 = 0.53$, $p = .766$	$X^2 = 3.26$, $p = .196$	$X^2 = 0.54$, $p = .764$
Would you share your favorite toy with Sila?	$X^2 = 0.27$, $p = .875$	$X^2 = 0.27$, $p = .875$	$X^2 = 0.26$, $p = .878$

References

Ahmad, M. I., Mubin, O., & Orlando, J. (2017). Adaptive social robot for sustaining social engagement during long-term children–robot interaction. *International Journal of Human-Computer Interaction, 33*(12), 943–962. https://doi.org/10.1080/10447318.2017.1300750.

Bastian, B., Jetten, J., & Ferris, L. J. (2014). Pain as social glue: Shared pain increases cooperation. *Psychological Science, 25*(11), 2079–2085. https://doi.org/10.1177/0956797614545886.

Beier, J. S., Gross, J. T., Brett, B. E., Stern, J. A., Martin, D. R., & Cassidy, J. (2019). Helping, sharing, and comforting in young children: Links to individual differences in attachment. *Child Development, 90*(2), e273–e289. https://doi.org/10.1111/cdev.13100.

Beneteau, E., Kientz, J. A., Richards, O. K., Yip, J., Zhang, M., & Hiniker, A. (2019). Communication breakdowns between families and Alexa. *2019 CHI Conference on Human Factors in Computing Systems Proceedings*. https://doi.org/10.1145/3290605.3300473.

Beran, T. N., Ramirez-Serrano, A., Kuzyk, R., Fior, M., & Nugent, S. (2011). Understanding how children understand robots: Perceived animism in child–robot interaction. *International Journal of Human-Computer Studies, 69*(7), 539–550. https://doi.org/10.1016/j.ijhcs.2011.04.003.

Biele, C., Jaskulska, A., Kopec, W., Kowalski, J., Skorupska, K., & Zdrodowska, A. (2019). How might voice assistants raise our children? In W. Karwowski, & T. Ahram (Eds.), *Intelligent human systems integration 2019* (pp. 162–167). Springer International Publishing.

Breazeal, C., Harris, P. L., DeSteno, D., Westlund, J. M. K., Dickens, L., & Jeong, S. (2016). Young children treat robots as informants. *Topics in Cognitive Science, 8*(2), 481–491. https://doi.org/10.1111/tops.12192.

Brielmann, A. A., & Stolarova, M. (2014). A new standardized stimulus set for studying need-of-help recognition (NeoHelp). *PloS One, 9*(1), Article e84373. https://doi.org/10.1371/journal.pone.0084373.

Chang, C.-W., Lee, J.-H., Chao, P.-Y., Wang, C.-Y., & Chen, G.-D. (2010). Exploring the possibility of using humanoid robots as instructional tools for teaching a second language in primary school. *International Forum of Educational Technology & Society, 13*(2), 13–24. https://doi.org/10.2307/jeductechsoci.13.2.13.

Common Sense Media. (2013). Zero to eight: Children's media sse in America 2013. https://www.commonsensemedia.org/research/zero-to-eight-childrens-media-use-in-america-2013.

Davies, C., & Katsos, N. (2010). Over-informative children: Production/comprehension asymmetry or tolerance to pragmatic violations? *Lingua, 120*(8), 1956–1972. https://doi.org/10.1016/j.lingua.2010.02.005.

Decety, J., Bartal, I. B.-A., Uzefovsky, F., & Knafo-Noam, A. (2016). Empathy as a driver of prosocial behaviour: Highly conserved neurobehavioural mechanisms across species. *Philosophical Transactions of the Royal Society B: Biological Sciences, 371* (1686), 20150077. https://doi.org/10.1098/rstb.2015.0077.

Egger, H. L., Pine, D. S., Nelson, E., Leibenluft, E., Ernst, M., Towbin, K. E., & Angold, A. (2011). The nimh child emotional faces picture set (NIMH-ChEFS): A new set of children's facial emotion stimuli. *International Journal of Methods in Psychiatric Research, 20*(3), 145–156. https://doi.org/10.1002/mpr.343.

Eisenberg, N., Miller, P. A., Shell, R., McNalley, S., & Shea, C. (1991). Prosocial development in adolescence: A longitudinal study. *Developmental Psychology, 27*(5), 849–857. https://doi.org/10.1037/0012-1649.27.5.849.

Eskritt, M., Whalen, J., & Lee, K. (2008). Preschoolers can recognize violations of the Gricean maxims. *British Journal of Developmental Psychology, 26*(3), 435–443. https://doi.org/10.1348/026151007X253260.

Fehr, E., Bernhard, H., & Rockenbach, B. (2008). Egalitarianism in young children. *Nature, 454*(7208), 1079–1083. https://doi.org/10.1038/nature07155.

Fehr, E., & Fischbacher, U. (2004). Social norms and human cooperation. *Trends in Cognitive Sciences, 8*(4), 185–190. https://doi.org/10.1016/j.tics.2004.02.007.

Genner, S., Suter, L., Waller, G., Schoch, P., Willemse, I., & Süss, D. (2017). Mike: Medien, Interaktion, Kinder, Eltern: Ergebnisbericht zur MIKE-Studie 2017. https://doi.org/10.21256/zhaw-3631.

Goodman, R. (1997). The strengths and difficulties questionnaire: A research note. *The Journal of Child Psychology and Psychiatry and Allied Disciplines, 38*(5), 581–586. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x. Scopus.

Grice, H. P. (1989). *Studies in the way of words*. Harvard University Press.

Gweon, H., Pelton, H., Konopka, J. A., & Schulz, L. E. (2014). Sins of omission: Children selectively explore when teachers are under-informative. *Cognition, 132*(3), 335–341. https://doi.org/10.1016/j.cognition.2014.04.013.

Hirschberg, J., & Manning, C. D. (2015). Advances in natural language processing. *Science, 349*(6245), 261–266. https://doi.org/10.1126/science.aaa8685.

Hoy, M. B. (2018). Alexa, Siri, Cortana, and more: An introduction to voice assistants. *Medical Reference Services Quarterly, 37*(1), 81–88. https://doi.org/10.1080/02763869.2018.1404391.

Isbister, K., & Nass, C. (2000). *Consistency of personality in interactive characters: Verbal cues, non-verbal cues, and user characteristics*.

Kahn, Peter H., Friedman, Batya, & Pérez-Granados, Deanne R. (2006). Robotic pets in the lives of preschool children. *Interaction Studies, 7*(3), 405–436. https://doi.org/10.1075/is.7.3.13kah.

Kahn, P. H., Gary, H. E., & Shen, S. (2013). Children's social relationships with current and near-future robots. *Child Development Perspectives, 7*(1), 32–37. https://doi.org/10.1111/cdep.12011.

Kahn, P. H., Jr., Kanda, T., Ishiguro, H., Freier, N. G., Severson, R. L., Gill, B. T., Ruckert, J. H., & Shen, S. (2012). "Robovie, you'll have to go into the closet now": Children's social and moral relationships with a humanoid robot. *Developmental Psychology, 48*(2), 303–314. https://doi.org/10.1037/a0027033.

Kennedy, J., Baxter, P., Senft, E., & Belpaeme, T. (2015). Higher nonverbal immediacy leads to greater learning gains in child-robot tutoring interactions. *Social Robotics*, 327–336.

Kim, E. H., Kwak, S. S., & Kwak, Y. K. (2009). Can robotic emotional expressions induce a human to empathize with a robot? RO-MAN 2009 - the 18th IEEE international symposium on robot and human interactive communication. https://doi.org/10.1109/ROMAN.2009.5326282, 358, 362.

Klasen, H., Woerner, W., Rothenberger, A., & Goodman, R. (2003). The German version of the Strengths and Difficulties Questionnaire (SDQ-Deu)—overview over first validation and normative studies. *Praxis der Kinderpsychologie und Kinderpsychiatrie, 52*(7), 491–502. Scopus.

Kohn, A. (1992). *No contest: The case against competition*. Houghton Mifflin Harcourt.

Lindenfors, P. (2017). For whose benefit?. *The biological and cultural evolution of human cooperation*. Springer.

Liu, C., Conn, K., Sarkar, N., & Stone, W. (2008). Online affect detection and robot behavior adaptation for intervention of children with autism. *IEEE Transactions on Robotics, 24*(4), 883–896. https://doi.org/10.1109/TRO.2008.2001362.

Lovato, S., & Piper, A. M. (2015). "Siri, is this you?": Understanding young children's interactions with voice input systems. *Proceedings of the 14th International Conference on Interaction Design and Children*, 335–338. https://doi.org/10.1145/2771839.2771910.

Mayer, R. E., Sobko, K., & Mautone, P. D. (2003). Social cues in multimedia learning: Role of speaker's voice. *Journal of Educational Psychology, 95*(2), 419–425. https://doi.org/10.1037/0022-0663.95.2.419.

McReynolds, E., Hubbard, S., Lau, T., Saraf, A., Cakmak, M., & Roesner, F. (2017). Toys that listen: A study of parents, children, and internet-connected toys. In *Proceedings of the 2017 CHI conference on human factors in computing systems* (pp. 5197–5207). https://doi.org/10.1145/3025453.3025735.

Melson, G. F., Kahn, P. H., Beck, A., Friedman, B., Roberts, T., Garrett, E., & Gill, B. T. (2009). Children's behavior toward and understanding of robotic and living dogs. *Journal of Applied Developmental Psychology, 30*(2), 92–102. https://doi.org/10.1016/j.appdev.2008.10.011.

Moore, C. (2009). Fairness in children's resource allocation depends on the recipient. *Psychological Science, 20*(8), 944–948. https://doi.org/10.1111/j.1467-9280.2009.02378.x.

Moreno, R., & Mayer, R. E. (2000). Engaging students in active learning: The case for personalized multimedia messages. *Journal of Educational Psychology, 92*(4), 724–733. https://doi.org/10.1037/0022-0663.92.4.724.

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues, 56*(1), 81–103. https://doi.org/10.1111/0022-4537.00153.

Pavey, L., Greitemeyer, T., & Sparks, P. (2011). Highlighting relatedness promotes prosocial motives and behavior. *Personality and Social Psychology Bulletin, 37*(7), 905–917. https://doi.org/10.1177/0146167211405994.

Pelikan, H. R. M., & Broth, M. (2016). Why that Nao?: How humans adapt to a conventional humanoid robot in taking turns-at-talk. *Proceedings of the 2016 CHI conference on human factors in computing systems*. https://doi.org/10.1145/2858036.2858478.

Reeves, B., & Nass, C. I. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. Cambridge University Press.

Sandoval, E. B., Brandstetter, J., Obaid, M., & Bartneck, C. (2016). Reciprocity in human-robot interaction: A quantitative approach through the prisoner's dilemma and the ultimatum game. *International Journal of Social Robotics, 8*(2), 303–317. https://doi.org/10.1007/s12369-015-0323-x.

Shahid, S., Krahmer, E., & Swerts, M. (2014). Child–robot interaction across cultures: How does playing a game with a social robot compare to playing a game alone or with a friend? *Computers in Human Behavior, 40*, 86–100. https://doi.org/10.1016/j.chb.2014.07.043.

Sinoo, C., van der Pal, S., Blanson Henkemans, O. A., Keizer, A., Bierman, B. P., Looije, R., & Neerincx, M. A. (2018). Friendship with a robot: Children's perception of similarity between a robot's physical and virtual embodiment that supports diabetes self-management. *Patient Education and Counseling, 101*(7), 1248–1255. https://doi.org/10.1016/j.pec.2018.02.008.

Sommer, K., Nielsen, M., Draheim, M., Redshaw, J., Vanman, E. J., & Wilks, M. (2019). Children's perceptions of the moral worth of live agents, robots, and inanimate objects. *Journal of Experimental Child Psychology, 187*, 104656. https://doi.org/10.1016/j.jecp.2019.06.009.

Sproull, L., Subramani, M., Kiesler, S., Walker, J. H., & Waters, K. (1996). When the interface is a face. *Human-Computer Interaction, 11*(2), 97–124. https://doi.org/10.1207/s15327051hci1102_1.

Tanaka, F., Cicourel, A., & Movellan, J. R. (2007). Socialization between toddlers and robots at an early childhood education center. *Proceedings of the National Academy of Sciences, 104*(46), 17954–17958. https://doi.org/10.1073/pnas.0707769104.

Tomasello, M. (2008). *Origins of human communication*. MIT.

Tomasello, M., Melis, A. P., Tennie, C., Wyman, E., & Herrmann, E. (2012). Two key steps in the evolution of human cooperation: The interdependence hypothesis. *Current Anthropology, 53*(6), 673–692. https://doi.org/10.1086/668207. JSTOR.

Turkle, S. (2005). *The second self: Computers and the human spirit*. MIT Press.

Vollmer, A.-L., Read, R., Trippas, D., & Belpaeme, T. (2018). Children conform, adults resist: A robot group induced peer pressure on normative social conformity. *Science Robotics, 3*(21). https://doi.org/10.1126/scirobotics.aat7111.

Warneken, F., Hare, B., Melis, A. P., Hanus, D., & Tomasello, M. (2007). Spontaneous altruism by chimpanzees and young children. *PLoS Biology, 5*(7), e184. https://doi.org/10.1371/journal.pbio.0050184.

Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science, 311*(5765), 1301–1303. https://doi.org/10.1126/science.1121448.

Webster, F. (2002). *Theories of the information society*. Routledge.

Westlund, J. M. K., Dickens, L., Jeong, S., Harris, P. L., DeSteno, D., & Breazeal, C. L. (2017). Children use non-verbal cues to learn new words from robots as well as people. *International Journal of Child-Computer Interaction, 13*, 1–9. https://doi.org/10.1016/j.ijcci.2017.04.001.

Williams, A., O'Driscoll, K., & Moore, C. (2014). The influence of empathic concern on prosocial behavior in children. *Frontiers in Psychology, 5*. https://doi.org/10.3389/fpsyg.2014.00425.

Yarosh, S., Thompson, S., Watson, K., Chase, A., Senthilkumar, A., Yuan, Y., & Brush, A. J. B. (2018). Children asking questions: Speech interface reformulations and personification preferences. *Proceedings of the 17th ACM Conference on Interaction Design and Children*, 300–312. https://doi.org/10.1145/3202185.3202207.